EFFECTS OF PHONOLOGICAL NEIGHBORHOOD DENSITY ON WORD PRODUCTION IN KOREAN

Jeffrey J. Holliday[a], Rory Turnbull[b]

[a]Indiana University, USA; [b]Ohio State University, USA
jeffh@ling.ohio-state.edu; turnbull@ling.ohio-state.edu

ABSTRACT

Studies have offered conflicting evidence for the effects of phonological neighborhood density (PND) on word production. Firm conclusions have been further hindered by the fact that most research on the effects of PND has been conducted on English. The purpose of this study was to provide an additional perspective by examining the effects of PND on word production in Korean. It was found that although PND correlated with both word duration and vowel duration, the effects were substantially diminished once certain segmental influences were controlled for. Additionally, no effect of PND was found on the VOT or f0 accompanying word-initial lax obstruents. We conclude that understanding variation in segmental content is vital to understanding effects of PND.

Keywords: Korean, neighborhood density, frequency, duration, VOT.

1. INTRODUCTION

A growing body of research into how phonological neighborhood density (PND) influences speech production has provided a variety of empirical results and new ways of thinking about the organization of the lexicon. The general trend observed in the literature is that vowels in words in denser neighborhoods tend to be more peripheral in the vowel space [8, 19, 20, 25, 26] than vowels in words in sparser neighborhoods.

The results for effects on duration, however, are far less clear, with limited evidence for higher PND correlating with longer word [17] and vowel durations [4, 5, 22], and more extensive evidence *against* the existence of such effects for word [12] and vowel durations [20, 23]. However, evidence exists in the consonantal domain in favor of cue enhancement, such as VOT lengthening in voiceless stops in high-density words [2, 9, 13, 16, 21].

On the other hand, effects of lexical frequency on duration are well-established, with more frequent words tending to have shorter word and vowel durations [5, 4, 10, 14, 24] than less frequent words. Much work on neighborhood density has attempted to control for frequency effects, either by including frequency as a covariate, or explicitly taking frequency into consideration in the density computation [5, 4, 20, 22, 23, 26].

These effects are purported to be language-general consequences of cognitive architecture or universal tendencies, but some of the evidence is contradictory [11, 12, 24] and the vast majority of work has been done on English. The purpose of the present study was to examine whether neighborhood density effects obtain in Korean, a language both typologically and genetically distinct from English.

In contrast with English, Korean is not stress-timed and lacks stress entirely, meaning that variation in duration between syllables is minimal [1, 7]. Additionally, Korean has a three-way phonation type contrast in word-initial stops and affricates cued jointly by voice onset time (VOT) and the fundamental frequency (f0) of the following vowel. Most relevant for the current study, it has been shown [15] that Seoul Korean speakers highlight the f0 (but not VOT) difference between lax and aspirated stops in clear speech, such that hyperarticulated lax stops are produced with a lower f0 and aspirated stops with a higher f0. We therefore hypothesized that, in light of analogous findings in English [2], that initial lax stops in high-density words in Korean would be pronounced with an enhanced f0, but not VOT, relative to those in low-density words.

2. METHODS

2.1. Materials

365 bisyllabic Korean words were selected to sample across a wide range of both PND and lexical frequency. Frequency information came from the Frequency Survey of Modern Korean Usage [18], a balanced corpus compiled from a variety of genres by the National Institute of the Korean Language, containing 1,531,966 word tokens and 58,437 word types. PND was calculated for each entry in the

KNC as the number of words with a phoneme edit distance of one from the current entry; that is, differed by one phoneme through insertion, deletion, or substitution.

All selected stimuli were of the form CVCVC or CVCCVC. The onset and medial consonants were taken from the set of legal onset consonants in Korean, excluding /l/ and /h/, and the word-final codas were taken from the set of legal coda consonants /p, t, k, m, n, ŋ/, excluding /l/. The lexical frequency of the stimuli ranged from 1 to 1,799 (median: 44), and PND ranged from 1 to 51 (median: 8).

2.2. Participants and procedure

Eight native speakers of Seoul Korean (4 male, 4 female) were recorded. Participants read the words one at a time in isolation, either from a printed list, or from a computer screen with each word presented individually.

2.3. Analysis

For each word token, word duration and duration of the second vowel (V2) were measured in milliseconds. The first vowel was not measured due to phonological processes in Korean which can cause the initial vowel to devoice or nasalize, rendering duration difficult to measure. For tokens with word-initial lax stops and affricates, VOT of the initial obstruent (C1) was measured in milliseconds, as was the mean f0, in ERB, over the first 25ms of the first vowel (V1).

We expected the number of phonemes in the word to have an effect on word duration, as it is uncontroversial that words with more phonemes tend to be longer than words with fewer phonemes. Since the stimuli in the current study involved both words with 5 phonemes (CVCVC) and 6 phonemes (CVCCVC), the modeling procedure controlled for the number of phonemes in the target word. It is also generally accepted that individual segments can differ in terms of duration. Because coda stops are unreleased in Korean, we expected to find that coda stops (/p, t, k/) would have a smaller contribution to word duration than coda nasals (/m, n, ŋ/). Therefore, the modeling procedure also controlled for the number of coda stops in a word. Note here that the number of coda stops and the number of phonemes in the current stimuli are necessarily collinear: 6-phoneme words can have a word-medial coda stop (e.g. [kɨkt*an] 'extreme'), a word-final coda stop (e.g. [kaŋmak], 'cornea'), both (e.g. [mɛkp*ak], 'pulse'), or none (e.g. [nonmun] 'thesis'); while 5-phoneme words can either have one coda stop (e.g. [patak] 'floor') or none (e.g. [somun] 'rumor'), due to the resyllabification of the word-medial consonant in a CVCVC word as the onset of the second syllable.

Word duration, V2 duration, C1 VOT, and V1 f0 were modeled using linear mixed effects modeling [3]. The modeling procedure followed a basic step-up algorithm, where fixed effects were added to a basic model, at each step using log-likelihood testing to establish whether the additional effect contributed significantly to data likelihood. The basic model was an intercept-only model, with random intercepts for talker and word identity. The vowel duration model began with an additional fixed effect of vowel height (1=low, 2=mid, 3=high), since height is known to influence vowel duration in Korean [6].

Fixed effects were added in the following order: number of phonemes; log lexical frequency; PND; number of coda stops. For each new fixed effect added, a corresponding slope was also entered into the random effect for talker. If a fixed effect did not contribute significantly to data likelihood, it was not included in subsequent modeling. Once a fixed effect was established in the model, 2-way interaction terms between it and all other fixed effects present were also tested, and retained if significant. Interaction terms were not entered as random slopes due to the danger of model overfitting. All variables were mean-centered before being entered into the model.

The 365 word productions from the eight talkers yielded a total of 2,920 tokens. Eleven tokens were excluded due to repetition, mispronunciation, or noise in the recording, providing a total of 2,909 tokens entered into the analysis, 1,600 of which had initial lax stops or affricates. 29 lax-initial tokens did not have a measurable f0 due to vowel devoicing, and thus were excluded from the f0 analysis.

3. RESULTS

3.1. Word duration

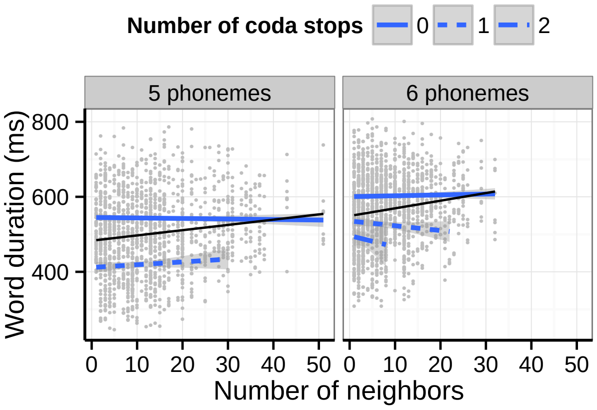
In modeling word duration, the addition of number of phonemes significantly contributed to data likelihood ($\chi^2(3) = 194.8$, $p < .0001$), although log lexical frequency did not. The contribution of PND was significant ($\chi^2(4) = 28.7$, $p < .0001$), but the interaction between PND and number of phonemes was not. Number of coda stops was additionally found to significantly contribute to data likelihood ($\chi^2(5) = 487.8$, $p < .0001$). The interaction between number of coda stops and number of phonemes was also significant ($\chi^2(1) = 44.5$, $p < .0001$), but the interaction between number of coda stops and PND was not. The model output is summarized in Table 1.

Table 1: Output from the word duration model.

Fixed effect	β	SE	t
Intercept	-116.22	27.49	-4.23
Nphonemes	115.11	10.63	10.83
PND	-0.05	0.33	-0.17
Ncodastops	-124.15	11.48	-10.82
Nphon. $\times$ Ncod.	57.93	8.42	6.88

Fig. 1 illustrates the effects on word duration revealed by the model. The effect of number of phonemes was as predicted, with words with more phonemes being longer than words with fewer phonemes. Similarly, words with more coda stops were shorter than words with fewer coda stops, presumably due to the negligible contribution of Korean's unreleased coda stops to overall word duration. An interaction between these two effects was also observed, although it must be interpreted carefully, as the number of phonemes is collinear with number of coda stops, in that all words with two coda stops had 6 phonemes. Specifically, the duration difference between words with zero and words with one coda stop was greater than the duration difference between words with one and words with two coda stops.

Figure 1: The relationship between PND and word duration with separate regression lines for words with 0, 1, and 2 coda stops. The overall linear relationship between PND and word duration, ignoring number of coda stops, is indicated by the thin solid line.



To summarize the output of the model, given in Table 1, the effects of the number of phonemes and the number of coda stops far outweighed the effect of PND. For example, the predicted difference in duration between 5- and 6-phoneme words was 115 ms, and the predicted difference between 5-phoneme words with and without a coda stop was 124 ms. On the other hand, with a coefficient for PND of -0.05, the predicted effect of having, for example, 10 additional neighbors was a decrease in duration of only 0.5 ms.

The individual lines for each number of coda stops in Fig. 1 also illustrate that words with more coda stops tend to have fewer neighbors, which explains why the effect size of PND in the model was so small. That is, when the number of coda stops is ignored, as the thin solid lines show in Fig. 1, the relationship between PND and word duration seems much stronger, and suggests that a higher PND leads to longer duration. But because the Korean lexicon is such that words with coda stops happen to have fewer neighbors, and it also happens to be that Korean coda stops are unreleased and thus do not contribute as much to total word duration, failing to account for segmental information could lead to the spurious conclusion that PND has a large and significant effect on word duration in Korean.

3.2. Vowel duration

In the vowel duration model, vowel height contributed significantly to model fit ($\chi^2(3) = 212.2$, $p < .0001$), but the number of phonemes and log lexical frequency did not. A significant effect of PND was observed ($\chi^2(5) = 19.5$, $p < .001$), but not of its interaction with vowel height. Finally, the number of coda stops was not found to contribute significantly to model fit. The final model output is summarized in Table 2, and the effects of vowel height and PND visualized in Fig. 2.

The effect of vowel height revealed in the model was as expected, such that lower vowels were longer than higher vowels, as was that of PND, such that vowels in words in denser neighborhoods were longer than vowels in words in sparser neighborhoods. It should be noted that the effect of PND was not significant within the model itself ($t < 2$), but that the inclusion of this factor contributed to model fit, and thus the effect of PND should be interpreted with caution. The predicted increase in vowel duration given, for example, 10 additional neighbors was only 1.4 ms.

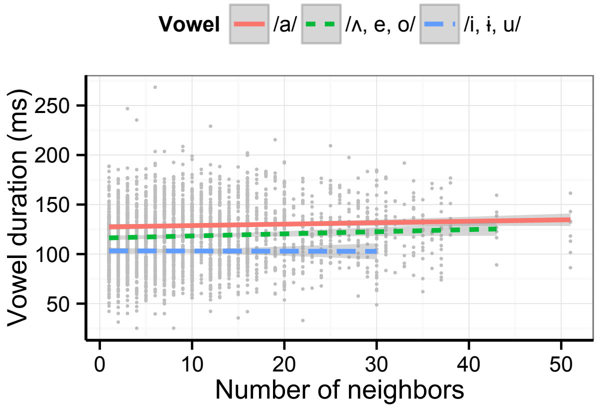
3.3. VOT and fundamental frequency

The final model for VOT contained only the number of phonemes, which was a significantly better model fit than the intercepts-only model ($\chi^2(3) = 23.1$, $p < .0001$). The model output is summarized in Table 3. The coefficient for the number of phonemes

Table 2: Output from the mixed effects model for V2 duration.

Fixed effect	β	SE	t
Intercept	24.75	6.15	4.03
Vowel Height	−12.63	1.56	−8.11
PND	0.14	0.12	1.13

Figure 2: The relationship between V2 duration and neighborhood density, split by vowel height.



in the final model predicts that the VOT of a word-initial lax stop in a 6-phoneme word is approximately 8.9 ms shorter than in a 5-phoneme word, and is unaffected by PND.

For f0, only the number of phonemes ($\chi^2(3) = 23.5$, $p < .0001$) was retained in the model. The model output is summarized in Table 4. The significant main effect of the number of phonemes was unexpected. Examination of the random slopes showed this effect was driven by one talker for whom the predicted difference between 5- and 6-phoneme words was approximately 5.7 Hz. For the remaining five talkers, the predicted absolute differences ranged from approximately 0.02 to 2.5 Hz, indicating that for most talkers the effect of the number of phonemes on f0 was very small.

The lack of an effect of PND on VOT is consistent with [15], who found that young speakers of Seoul Korean did not enhance VOT targets in their clear speech stop productions. The lack of an effect of PND on f0, however, was unexpected and is inconsistent with other data suggesting that high PND leads to cue enhancement in English stop production [2, 9, 13, 16, 21]. It is noteworthy, however, that a model with PND alone as a fixed effect showed a significant improvement to data likelihood over an intercept-only model ($\chi^2(3) = 8.0$, $p = .046$), although the observed effect was small in absolute terms: the difference between the mean f0s of the upper and lower 50% of words when split according to PND was 0.04 ERB (or 1.66 Hz). This possibility of an effect would need further investigation to clarify; it also highlights the importance of the order of variable selection in stepwise regression methods.

Table 3: Output from the mixed effects model for VOT of word-initial lax stops and affricates.

Fixed effect	β	SE	t
Intercept	−0.94	3.85	−0.24
Nphonemes	−8.87	2.65	−3.35

Table 4: Output from the mixed effects model for f0 of word-initial lax stops and affricates.

Fixed effect	β	SE	t
Intercept	−0.28	0.40	−0.71
Nphonemes	0.02	0.02	0.74

4. DISCUSSION AND CONCLUSION

The results of the current study provide mixed support for effects of PND on word production in Korean. Although PND was found to be necessary in the model-construction process for word and vowel duration, most of the variance it accounted for was more adeptly explained by segmental factors. For example, it was found that PND tended to be lower in words with more coda stops, which are presumably shorter in duration due to the stops' being unreleased, spuriously suggesting that a low PND leads to shorter word duration. This finding is consistent with the notion that the effects of PND on vowel space expansion can be similarly explained by segmental influences alone [11]. We conclude that neighborhood density may be subordinate to segmental content in the context of speech production in Korean.

5. ACKNOWLEDGEMENTS

Both authors contributed equally to this work. We are grateful to Jieun Kang and Woo-bong Shin of Korea University for assisting with data collection.

6. REFERENCES

[1] Arvaniti, A. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40, 351–373.

[2] Baese-Berk, M., Goldrick, M. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24, 527–554.

[3] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

[4] Burdin, R. S., Turnbull, R., Clopper, C. G. 2014. Interactions among lexical and discourse characteristics in vowel production. *Journal of the Acoustical Society of America* 136, 2172.

[5] Burdin, R. S., Turnbull, R., Clopper, C. G. 2014. Variability in vowel production: Exploring interactions among frequency, neighborhood density, predictability, and mention. *Journal of the Acoustical Society of America* 135, 2291.

[6] Chung, H., Kim, K., Huckvale, M. 1999. Consonantal and prosodic influences on Korean vowel duration. *Proceedings of Eurospeech*.

[7] Chung, Y., Arvaniti, A. 2013. Speech rhythm in Korean: Experiments in speech cycling. *Proceedings of Meetings on Acoustics* 19, 060216.

[8] Clopper, C. G., Tamati, T. N. 2014. Effects of local lexical competition and regional dialect on vowel production. *Journal of the Acoustical Society of America* 136(1), 1–4.

[9] Fox, N. P., Reilly, M., Blumstein, S. E. to appear. Phonological neighborhood competition affects spoken word production irrespective of sentential context. *Journal of Memory and Language*.

[10] Gahl, S. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84, 474–496.

[11] Gahl, S. 2015. Lexical competition in vowel articulation revisited: Vowel dispersion in the Easy/Hard database. *Journal of Phonetics* 49, 96–116.

[12] Gahl, S., Yao, Y., Johnson, K. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66, 789–806.

[13] Goldrick, M., Vaughn, C., Murphy, A. 2013. The effects of lexical neighbors on stop consonant articulation. *Journal of the Acoustical Society of America* 134(2), EL172–EL177.

[14] Jurafsky, D., Bell, A., Gregory, M., Raymond, W. D. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In: Bybee, J., Hopper, P., (eds), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins 229–254.

[15] Kang, K.-H., Guion, S. G. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America* 124, 3909–3917.

[16] Kirov, C., Wilson, C. 2012. The specificity of online variation in speech production. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* 587–592.

[17] Kryuchkova, T. V., Tucker, B. V. 2012. Emotion coexists with lexical effects: A case study. *Canadian Acoustics* 40(3), 32–33.

[18] The National Institute of the Korean Language, 2003. Kwuklipkwukeyenkwuwen sosik. *Saykwuke Saynghwal* 13(1), 185–208.

[19] Munson, B. 2013. The influence of production latencies and phonological neighborhood density on vowel dispersion. *Proceedings of Meetings on Acoustics* 19, 060192.

[20] Munson, B., Solomon, N. P. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research* 47, 1048–1058.

[21] Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., Baese-Berk, M. 2011. Phonological neighborhood effects in spoken word production: An fMRI study. *Journal of Cognitive Neuroscience* 23(3), 593–603.

[22] Scarborough, R. 2010. Lexical and contextual predictability: Confluent effects on the production of vowels. In: Fougeron, C., Kühnert, B., D'Imperio, M., Vallée, N., (eds), *Laboratory Phonology 10*. Berlin: Mouton de Gruyter.

[23] Scarborough, R. 2013. Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics* 41, 491–508.

[24] Turnbull, R. 2015. *Assessing the listener-oriented account of predictability-based phonetic reduction*. PhD thesis Ohio State University Columbus, OH.

[25] Watson, P. J., Munson, B. 2007. A comparison of vowel acoustics between older and younger adults. *ICPhS XVI*.

[26] Wright, R. 2004. Factors of lexical competition in vowel articulation. In: Local, J., Ogden, R., (eds), *Papers in Laboratory Phonology 6*. Cambridge: Cambridge University Press 26–50.